Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

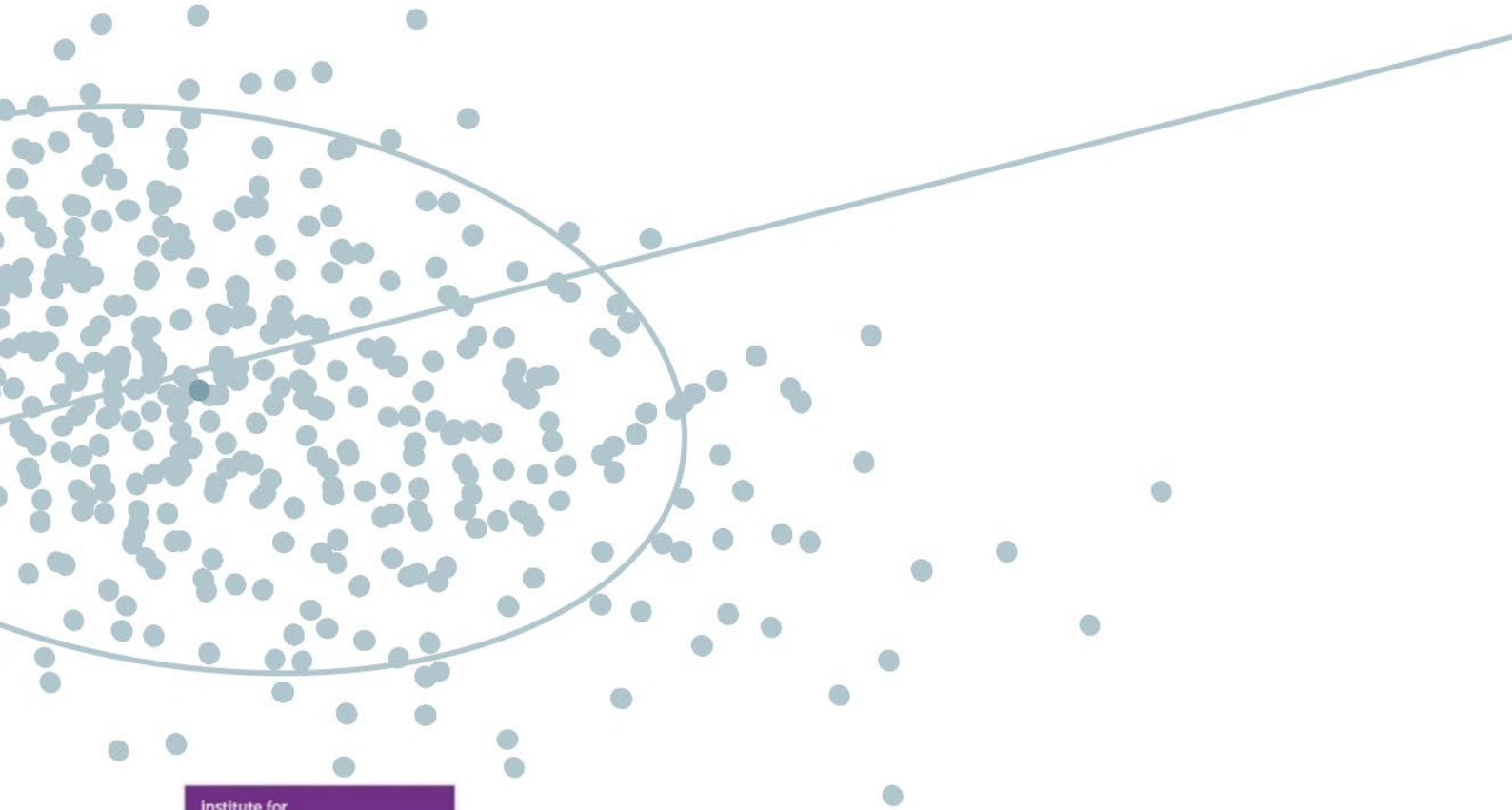
Federal Department of Home Affairs FDHA
Federal Office of Public Health FOPH

Health Technology Assessment (HTA)

HTA Protocol

Multigene-expression tests in early breast cancer

Version 30.04.2024



institute for
Medical
Technology
Assessment

Title	Multigene-expression tests in early breast cancer
Author/Affiliation	Gimon de Graaf, Stavros Anagnostopoulos, Martine Hoogendoorn-Lips Institute for Medical Technology Assessment, Erasmus University Rotterdam Anouk Oordt Epi-C (Epidemiology-Consultancy)
Technology	Diagnostic tests
Date	30 April 2024

Conflict of Interest:

The authors have no financial, academic, personal or any other conflicts of interest to declare in relation to this project.

Federal Office of Public Health FOPH
Health Technology Assessment
Schwarzenburgstrasse 157
CH-3003 Bern
Switzerland
Tel.: +41 58 462 92 30
E-mail: hta@bag.admin.ch

Cover image: Simplified representation of a fictional probabilistic sensitivity analysis on a cost-effectiveness plane with the diagonal willingness-to-pay line.

Executive Summary

Breast cancer is the most frequently diagnosed cancer among women. Treatment of early-stage breast cancer primarily involves a combination of local modalities (surgery, radiotherapy), systemic anticancer treatments (chemotherapy, endocrine therapy, molecularly targeted therapies) and other supportive measures. There are different subtypes of breast cancer. For the subtype oestrogen receptor positive (ER+), human epidermal growth factor receptor 2 negative (HER2-) breast cancers, more than 90% of patients are diagnosed in early non-metastatic stages. About 15% of these patients will have a recurrence within 5 years after initiating endocrine therapy. This recurrence risk can be reduced through the use of adjuvant chemotherapy. However, due to the adverse effects associated with this treatment, a decision must be made whether the benefits outweigh the risks for a particular patient. Clinical assessment, imaging techniques, and histopathological analysis of tumour tissue typically guide clinicians to select patients eligible for adjuvant chemotherapy. In addition to these conventional test modalities, multigene-expression tests have been developed to facilitate the selection.

Multigene-expression tests are temporarily covered by the mandatory health insurance in Switzerland since 2015 for patients with ER+/HER2- breast cancer with up to 3 affected lymph nodes, for whom the results of conventional testing alone do not allow a clear decision to be made regarding adjuvant chemotherapy. This HTA will support decision making regarding cost coverage of the in Switzerland approved multigene-expression tests (i.e. Oncotype DX, MammaPrint, EndoPredict and Prosigna). In this HTA protocol, a research question and the key questions covering the HTA domains of clinical efficacy/effectiveness/safety, budget impact/cost-effectiveness, and ethical/legal/social/organisational issues are formulated and the methodological approach to conduct the HTA is described. The HTA report following this HTA protocol aims to answer these questions based on the best available scientific evidence for the multigene-expression tests Oncotype DX, MammaPrint, EndoPredict and Prosigna for guiding adjuvant chemotherapy decisions in patients with ER+, HER2-, LN0-3 early breast cancer after surgical resection.

For the clinical evaluation, a systematic literature search of the PubMed (MEDLINE), Embase.com and Cochrane Library databases will be conducted covering the last 15 years. Search strings will be compiled for the broad population (breast cancer) and interventions (Oncotype DX, MammaPrint, EndoPredict and Prosigna). In case less than one prospective RCT is found for a multigene-expression test, an additional systematic literature search for retrospective/re-analyses of RCTs and comparative non-randomised studies will be conducted. Outcomes of interest are overall survival, disease-free survival, recurrence-free-survival, and health-related quality of life. In addition, it will be searched for before-and-after studies on the outcome impact on treatment management. The methodological quality of included studies will be critically appraised, and data is extracted. Pooled estimates will be calculated by meta-analysis for outcomes, provided the outcome measures can be combined. Outcomes for which it is not possible to calculate

pooled estimates will be analysed narratively. The overall certainty of the evidence on outcome level will be assessed with GRADE.

For the economic evaluation, the systematic literature search will follow the principles of the clinical evaluation. The search will encompass the databases PubMed (MEDLINE), Embase.com, Cochrane Library, Cost Effectiveness Analysis (CEA) registry, Tufts Medical Centre Cost-Effectiveness Analysis Registry, and National Health Service Economic Evaluation Database (NHS EDD) (last update on the 31st of March of 2015) to identify existing economic studies that are directly applicable to the research question. If the systematic literature search and study selection yield sufficient data applicable to the Swiss setting, an economic model will be developed that will utilise up-to-date Swiss-specific cost and clinical inputs, or alternatives that are most applicable to the Swiss context. Finally, a budget impact analysis will be conducted.

For the evaluation of ethical, legal, social and organisational domains, relevant issues will be identified from the studies included in the clinical evaluation. In addition, targeted non-systematic searches in PubMed (MEDLINE) and the grey literature will be conducted to identify information related to these domains; key issues will be summarised narratively.

Table of Contents

Abbreviations and acronyms

CEA	Cost-effectiveness analysis
CHEC	Consensus Health Economic Criteria
CHEERS	Consolidated Health Economic Evaluation Reporting Standards
CHF	Swiss franc
CMA	Comprehensive Meta-analysis
EAE	Effectiveness, appropriateness and economic efficiency
EBCTCG	Early Breast Cancer Trialists' Collaborative Group
ELSO	Ethical, legal, social and organisational domains
ER	Oestrogen receptor
ESMO	European Society for Medical Oncology
EUnetHTA	European Network for Health Technology Assessment
FFPE	Formalin-Fixed Paraffin-Embedded
FOPH	Federal Office of Public Health
GRADE	Grading of Recommendations Assessment, Development and Evaluation
HER2	Human epidermal growth factor receptor 2
HR	Hormone receptor
HRQoL	Health-related quality of life
HTA	Health Technology Assessment
ICER	Incremental cost-effectiveness ratio
LN	Lymph nodes
PICO	Population, intervention, comparator, outcome
PR	Progesterone receptor
PRISMA	Preferred Items for Systematic Reviews and Meta-Analyses
QALY	Quality-adjusted life year
qRT-PCR	Quantitative reverse transcription polymerase chain reaction

RCT	Randomised controlled trial
RNA	Ribonucleic acid
RoB 2	Revised Cochrane Risk of Bias tool for randomised trials
ROBINS-I	Risk of Bias in Non-randomised Studies - of Interventions
ROR	Risk of Recurrence

Objective of the HTA Protocol

Based on a preliminary screening of the literature the objective of the health technology assessment (HTA) protocol is to formulate the HTA key questions, to define the population, intervention, comparator, outcomes (PICO) and describe the methodology to conduct a systematic literature search, extract, analyse and synthesise the data in the HTA report on the topic. Key questions are defined, addressing the main HTA domains, i.e. clinical efficacy/effectiveness/safety, budget impact/cost-effectiveness, ethical/legal/social and organisational issues.

1. Policy question and context

Each health technology assessment (HTA) topic entails a policy and a research question. In healthcare, a policy question is a request to regulate a reimbursement policy and is aimed at securing financing of health technologies. Such a request, related to a particular health technology, typically addresses an existing controversy around a technology.

The HTA Unit of the Federal Office of Public Health (FOPH) was mandated to re-evaluate the available evidence for the use of multigene-expression tests in early breast cancer patients until summer 2025. The policy question was defined as:

> *Do the multigene-expression tests Oncotype DX, MammaPrint, EndoPredict and Prosigna meet the effectiveness, appropriateness and economic efficiency (EAE) criteria to guide decision making on adjuvant chemotherapy in patients with oestrogen receptor positive (ER+), human epidermal growth factor receptor 2 negative (HER2-) and up to 3 positive lymph nodes (LN0-3) early breast cancer after surgical resection for whom it is unclear based on conventional testing whether to prescribe adjuvant chemotherapy?*

The policy question refers to an "unclear" population based on conventional testing whether to prescribe adjuvant chemotherapy. This group of patients is described in the current reimbursement text for cost coverage by the Swiss compulsory health insurance as a population for whom conventional test findings alone do not allow a clear decision to be made regarding adjuvant chemotherapy (original text in German see: *Anhang 1 der Krankenpflege-Leistungsverordnung [KLV]*)[1].

Furthermore, this "unclear" population was confirmed as the typical target population for the use of multigene-expression tests in Switzerland by an advisory expert group consisting of 6 Swiss clinical experts in the field of breast cancer treatment and diagnosis at a meeting held by the FOPH in 2023 (see ***Appendix A***). However, during the same meeting, it became evident that the identification of the target population (i.e. unclear based on conventional testing) is only possible to a certain extend and is rather done by excluding patients that can be classified low or high risk for disease recurrence based on conventional testing.

The definition of the target population in the reimbursement text is also in line with the panel recommendations for genomic signature testing (i.e. multigene-expression tests) of the St Gallen International Breast Cancer Consensus Conferences 2021 and 2023.[1,2] At the international conference in 2023 the target population was defined as "ER+, HER2- early breast cancer where the indication for adjuvant chemotherapy is considered uncertain". At the conference in 2021, the majority of panellists favoured consideration of multigene-expression tests in select cases as opposed to routine testing in ER+, HER2-, LN0-3 early breast cancer. This means that physicians consider

[1] This can be found at https://www.bag.admin.ch/bag/de/home/versicherungen/krankenversicherung/krankenversicherung-leistungen-tarife/Aerztliche-Leistungen-in-der-Krankenversicherung/anhang1klv.html

themselves to be able to identify a relevant subgroup of patients for the application of multigene-expression tests in the absence of definitive selection criteria. The 2 conference reports did not define the subgroup more explicitly. Similarly, in the 2024 European Society for Medical Oncology (ESMO) guidelines for early breast cancer, in case of HR+, HER2- patients for whom there is "uncertainty about indications for adjuvant chemotherapy decision after consideration of all clinical and pathological factors", gene expression tests or endocrine response assessment are recommended to guide the treatment decision.[3]

This HTA will assess the availability of evidence for the target population(s) of the respective multigene-expression tests as intended by the manufacturers (with the exception of ER- and HER2+ patient populations). These target populations may or may not be identical to the population of the policy question. The report will identify and evaluate for which groups of patients there is evidence in the peer-reviewed literature of the individual multigene-expression tests Oncotype DX, MammaPrint, EndoPredict and Prosigna. If provided in the literature, the HTA will evaluate the evidence for the subgroup of patients whose conventional testing results were "unclear". If no evidence is found that directly matches the Swiss clinical setting, the best available surrogate evidence for the Swiss clinical setting will be identified. This is reflected in the research question presented in the following chapter as well as the HTA key questions in ***Chapter 6***.

2. Research Question

Based on the policy question and context described above, the following main research question was formulated:

> *What is the clinical effectiveness and cost-effectiveness of the multigene-expression tests Oncotype DX, MammaPrint, EndoPredict and Prosigna for guiding adjuvant chemotherapy decisions in patients with ER+, HER2-, LN0-3 early breast cancer after surgical resection?*

A particular interest for this research question is the "unclear" population, as described in the previous chapter. As the 4 different multigene-expression tests are not interchangeable, the research question considers each test individually and not collectively as a category of tests.

3. Medical background

Breast cancer is a common disease in Switzerland and the second most common cancer type in the country with approximately 6,300 women and 50 men being newly diagnosed every year.[4,5] Survival rates in the Swiss population are 87.9% (95% CI 87.3–88.5%) at 5 years and 80.1% (95% CI 79.2–81.0%) at 10 years.[6] The disease is caused by an uncontrolled reproducing of cells in the epithelia of the ducts or the lobules of the breast or the tissue in between due to genetic mutations, resulting in a tumour.[7] Breast cancer is a heterogeneous disease and is traditionally classified by

expression of the oestrogen receptor (ER), progesterone receptor (PR) and/or human epidermal growth factor receptor (HER) on the surface of the cancer cells.[8] Hormone receptor (HR) positive (i.e. ER+ and/or PR+), HER2- breast cancer represents about 70% of breast cancer diagnosis in Western countries.[9]

Breast cancer has only a few obvious symptoms. Especially, early breast cancers, including early HR+, HER2- breast cancers, are usually asymptomatic. Symptoms include breast thickenings or lumps, changes in the appearance of the breast (shape, size etc.), redness, dimpling, or pitting in the skin, changes in the nipple or surrounding area, and abnormal or bloody discharge from the nipple.[10]

The diagnosis of breast cancer relies on physical examination, imaging techniques, and histopathological analysis of the tumour tissue for the assessment of the extent and management of the disease. Clinical examination entails palpation and imaging tests including mammography, ultrasonography and MRI.[11] Pathologic evaluation entails immunohistochemical and molecular tests on tissue obtained by fine-needle aspiration, core biopsy or surgical excision.[11] Breast cancer staging is determined based on tumour size and location, lymph node involvement, metastasis to other body parts, tumour grade, and the presence of specific biomarkers.[12] Regarding newly diagnosed breast cancers, physical examination, mammography or ultrasound are often sufficient for locoregional staging, while staging for early breast cancers in clinical stages I and II can be achieved by routine blood tests.[11]

The treatment of early breast cancers includes a combination of local modalities (surgery, radiotherapy), systemic anticancer treatments (chemotherapy, endocrine therapy, molecularly targeted therapies) and other supportive measures.[13] The 2021 St. Gallen International Consensus Guidelines recommend that adjuvant systemic therapy (therapy following surgery) is considered for nearly all patients with early invasive breast cancer, depending on the evaluation of the prognostic and predictive factors and the potential benefits and side effects of the treatment.[1] More specifically, adjuvant endocrine therapy is recommended for all ER+ tumours for 5 to 10 years, while in cases where the individual risk of recurrence or the disease burden are high, adjuvant chemotherapy may be required.[1,13] The use of predictive biomarkers such as ER, PR, HER2, and Ki67 (nuclear protein being associated with cellular proliferation) and approved genomic signatures are established to help in forming treatment decisions.[13,14] Additionally, the patient's age, menopausal status, comorbidities, overall health status, and personal preferences are considered.[13]

In general, better prognosis is associated with small tumour size, lymph node-negative (LN0) status, younger age and ER+/PR+ status. Even though more than 90% of HR+/HER2- breast cancer primary diagnoses are in early non-metastatic stages, it is observed that approximately 15% of patients with node-positive HR+ HER2- breast cancer will have a recurrence within 5 years of initiating endocrine therapy, while the risk of a recurrence is higher when high-risk features (e.g. node-positive status) are present.[8,15]

Meta-analyses of randomised clinical trials (RCTs) by the Early Breast Cancer Trialists' Collaborative Group (EBCTCG) have showed that adjuvant chemotherapy (chemotherapy following surgery) is associated with a reduced risk of cancer recurrence and death in people with early-stage breast cancer, presumably by treating micro metastases that may not be clinically evident.[16,17] However, chemotherapy is correlated with short- or long-term side effects that result in additional costs and reduced quality of life. Therefore, the choice of initiating adjuvant chemotherapy should be based on a trade-off between the potential hams and benefits, taking into account the patient's individual risk of recurrence, predicted sensitivity to the treatment, age, comorbidities and preferences.[13]

4. Technology description

In the last 20 years, multigene-expression tests have been developed to guide the decision to initiate adjuvant chemotherapy by categorising patients with early breast cancer into different groups based on the risk of recurrence by analysing the expression of various genes.[18] Multiple multigene-expression tests have been developed that differ regarding employed techniques and the specific kind of genes that are measured, they however all examine genes related to factors such as proliferation or oestrogen receptor pathway, with proliferation genes having the biggest impact on the assessment of prognosis.[19] These tests have predominantly been developed and evaluated to predict the probability of recurrence (i.e. a prognostic test). However, some have also been evaluated for their ability to predict benefit of adjuvant chemotherapy (i.e. as a predictive test). In practice, multigene-expression tests are used to identify those patients for whom the survival and quality of life benefit from adjuvant chemotherapy.

There are 4 multigene-expression tests (Oncotype DX, MammaPrint, EndoPredict, Prosigna) currently covered by the mandatory health insurance in Switzerland, which are described in the following sections.

4.1 Oncotype DX

Oncotype DX is a multigene-expression test that assesses the probability of distant relapse within a 9-year span assuming 5 years of endocrine therapy and predicts the anticipated response to adjuvant chemotherapy. The intended population consists of pre- and post-menopausal patients with early-stage ER+/HER2- breast cancer with 0 to 3 positive nodes. The assay measures the expression of 21 genes (16 cancer-related genes and 5 reference genes) that were identified by correlating the expression of 250 genes with recurrence-free survival in 3 clinical trials, in order to generate a risk of recurrence score (RS) that ranges from 0 to 100.[20] Based on the results of the NSABP B20 trial, the relationship between the RS and the magnitude of the chemotherapy benefit was found statistically significant.[21] The measurement is performed on Formalin-Fixed Paraffin-Embedded (FFPE) tissue sample from a biopsy or surgical resection and it utilises the quantitative reverse transcription polymerase chain reaction (qRT-PCR) technology.[20] The RS has been

validated by the NSABP B-20 study and the TransATAC and SWOG 8814 trials, while the most recent cut-off points for the risk categories are derived from the TAILORx trial.[21–24] They consist of low (RS<11), intermediate (11≤RS<26), and high risk (RS≥26).[24,25]

4.2 MammaPrint

MammaPrint is a microarray-based multigene-expression test that measures the mRNA expression of 70 genes related to the 10 hallmarks of cancer namely: apoptosis, self-sufficiency in growth signals, insensitivity to anti-growth signals, limitless replicative potential, tissues invasion and metastasis, sustained angiogenesis, genome instability and mutation, tumour-promoting inflammation, deregulating cellular energetics, and avoiding immune distruction.[26] The test is intended for pre- and post-menopausal patients with early breast cancer with 0 or up to 3 positive lymph nodes irrespective of their ER/HER2 status. It assigns tumours into risk categories for distant recurrence within 5 and 10 years, while also assessing which patients are less likely to benefit from adjuvant chemotherapy.[27–30] The sample is processed by isolating mRNA from a FFPE sample and a score (MammaPrint Index) between -1 and 1 is calculated.[31,32] Samples with a MammaPrint index value greater than 0 are classified as low risk, and samples with a value less than or equal to 0 are classified as high risk. The test was validated in the MINDACT trial.[33,34] MammaPrint can be combined with an 80-gene molecular subtyping assay, called BluePrint, that categorises tumours in the luminal, HER2 or basal intrinsic subtypes.[35]

4.3 EndoPredict

EndoPredict is an assay intended for pre- or post-menopausal patients with ER+/HER2- breast cancer with node-negative or positive (up to 3 nodes) disease, that assesses the risk of 10-year distant recurrence (metastatic disease) from the time of initial diagnosis assuming 5 years of endocrine therapy, the likelihood of distant recurrence 5-15 years after diagnosis, and the estimated absolute benefit of chemotherapy at 10 year. This 12-gene expression test is based on the quantification of the RNA expression of 8 prognostic genes, 3 normalisation genes, and 1 control gene measured in FFPE resection and biopsy tissue samples by qRT-PCR.[36] These genes were selected out of 63 candidate genes and are linked to tumour proliferation and hormone receptor activity. A molecular score (EP) is generated that ranges between 0.0 and 15.0, and when combined with the clinical characteristics (tumour size and number of affected lymph nodes), the clinically applicable EPclin Score is derived that ranges between 1.0 and 8.2.[36–38] An EP score of <5 and an EPclin score of <3.3 are considered low risk for distant recurrence, while an EP score of ≥5 and EPclin score ≥3.3 are considered high risk.[36] The EP and EPclin scores were validated in the ABCSG-6, ABCSG-8 trials and TransATAC study.[36,39] Additionally, EndoPredict provides prognostic information on pathological factors such as Ki67.[36,39]

4.4 Prosigna

Prosigna is a test that estimates the risk of recurrence over a span of 10 years, assuming 5 years of endocrine therapy, for postmenopausal patients with ER+/HER2- early breast cancer with node-negative disease or up to 3 positive lymph nodes. The test examines the expression levels of 50 genes using the PAM50 gene panel utilising RNA expression extracted from an FFPE sample using the RNA hybridization technique and the nCounter System.[40] The assay measures the expression of 8 normalisation genes, 6 positive control genes and 8 negative control genes and classifies the tumours in one of the luminal A, luminal B, HER2 enriched and basal-like subtypes based on the results of the PAM50 results by generating a risk of recurrence score (ROR). The ROR-PT score that is utilised by Prosigna is derived from the assay when the subtype, tumour size and proliferation score are combined and its values range between 0 and 100. When accounting for nodal status, breast cancers are categorised in risk categories of low, intermediate or high and the corresponding scores are 0-40, 41-60, and 61-100 for node negative breast cancers and 0-15, 16-40, and 41-100 for node positive cancers (up to 3 nodes).[40] The ROR score was validated in the TransATAC, ABCSG-8 and DBCG cohorts. [41–43]

4.5 Overview of test characteristics

An overview of relevant characteristics of the 4 multigene-expression tests described in the previous sections is shown in ***Table 1*** below.

Table 1: Characteristics of the 4 multigene-expression tests covered in Switzerland

Commercial name	Oncotype DX	MammaPrint	EndoPredict	Prosigna
Applications indicated by manufacturer	Recurrence risk and chemotherapy benefit	Recurrence risk and chemotherapy benefit	Recurrence risk and chemotherapy benefit	Recurrence risk and intrinsic subtype
Technique	qRT-PCR	Microarray	qRT-PCR	RNA hybridization (NanoString nCounter system)
Genes	21-gene assay	70-gene assay	12-gene assay	50-gene assay
Hormone receptor status	ER+	ER+ or ER-	ER+	ER+
HER2 status	HER2-	HER2-/HER2+	HER2-	HER2-
Lymph node status	LN0 or LN+ (up to 3 positive nodes)	LN0 or LN+ (up to 3 positive nodes)	LN0 or LN+ (up to 3 positive nodes)	LN0 or LN+ (up to 3 positive nodes)
Menopausal status	Pre- and post-menopausal	Pre- and post-menopausal	Pre- and post-menopausal	Post-menopausal
Output	RS (0-100)	MPI (-1.000 to +1.000)	EP score (0.0-15.0) and EPclin score (1.0-8.2)	ROR score (0-100) and intrinsic subtype
Categorisation of output	Low (RS<11), intermediate (11≤RS<26), or high risk (RS≥26)	Low (MPI<0) or high risk (MPI>0)	Low (EP<5 & EPclin<3.3) or high risk (EP<5 & EPclin<3.3)	Low, intermediate, or high risk (categorisation conditional on nodal status, see main text)

Commercial name	Oncotype DX	MammaPrint	EndoPredict	Prosigna
Therapy assumed for prognosis	5 years of endocrine therapy	No therapy assumption	5 years of endocrine therapy	5 years of endocrine therapy
Manufacturer	Genomic Health	Agendia	Myriad Genetics	Veracyte

Abbreviations:
EP score = molecular score, ER = oestrogen receptor, HER2 = human epidermal growth factor receptor 2, LN= lymph node, MPI = MammaPrint Index, qRT-PCR = quantitative real-time reverse-transcription polymerase chain reaction, ROR = risk of recurrence, RS = recurrence score.

5. Population, Intervention, Comparator, Outcome (PICO)

Two PICOs are defined, aimed at covering a broad range of relevant outcomes and in order to employ different inclusion criteria for each. See ***Sections 7.1.1 and 7.1.2***. The first PICO focuses on the clinical and economic outcomes, whereas the second focuses on the impact on clinical decision making. Information from Swiss clinical experts gathered during an Advisory Clinical Expert Group meeting was used for the specification of these PICOs. A summary of the Advisory Clinical Expert Group meeting can be found in ***Appendix A***.

PICO 1 is defined as follows:

P:	Women and men with ER+, HER2-, LN0-3 early breast cancer[a] after surgical resection
I:	Conventional testing (with or without clinical prediction tools, such as Adjuvant!Online) and multigene-expression tests (Oncotype DX, MammaPrint, EndoPredict, or Prosigna)
C:	Conventional testing (with or without clinical prediction tools, such as Adjuvant!Online)
O:	**Clinical outcomes[a,b]:** - Overall survival - Disease-free survival - Recurrence-free-survival (i.e. invasive disease-free survival and distant recurrence-free survival) - Health-related quality of life **Economic outcomes** - Incremental/total costs, life years, and quality-of-life-adjusted life-years - Incremental cost-effectiveness ratio - Budget impact

Notes:
a = Data will be extracted as defined by the authors, including the definition of early breast cancer and the outcomes.
b = Depending on the design and comparator of the included studies, predictive ability (i.e. the degree to which the test predicts the chemotherapy benefits on clinical outcomes) and prognostic ability (i.e. the degree to which the test accurately determines the risk of a clinical outcome) of the multigene-expression tests will be assessed.

PICO 2 is defined as follows:

P:	Women and men with ER+, HER2-, LN0-3 early breast cancer[a] after surgical resection
I:	Conventional testing (with or without clinical prediction tools, such as Adjuvant!Online) and multigene-expression tests (Oncotype DX, MammaPrint, EndoPredict, or Prosigna)
C:	Conventional testing (with or without clinical prediction tools, such as Adjuvant!Online)
O:	Impact on treatment management (i.e. clinical impact according to the test result on the proportion of patients to receive adjuvant chemotherapy)

Notes:
a = Data will be extracted as defined by the authors, including the definition of early breast cancer.

6. HTA key questions

6.1 HTA research questions

For the evaluation of the technology the following research questions covering central HTA domains, as designated by the European Network for Health Technology Assessment (EUnetHTA) Core Model (clinical effectiveness, safety, cost-effectiveness, budget impact, ethical, legal, social and organisational aspects), are addressed:

1. Is the technology effective/efficacious compared to the comparator technology?
2. Is the technology safe compared to the comparator technology?
3. What is the budget impact of the technology?
4. Is the technology cost-effective compared to the comparator technology?
5. Are there ethical, legal, social or organisational issues related to the technology?

Note. Safety is not considered in the PICO, because multigene-expression tests are not expected to lead to any direct safety concerns. There are no direct adverse effects, because the tissue samples used for the multigene-expression tests origin from already surgically resected breast tissue. The indirect safety effects of multigene-expression tests, such as additional treatment that can cause harm compared to a situation without these tests, are incorporated within the survival and

HRQoL outcomes as defined in the PICO. In addition, safety will be addressed in the section of the report considering ethical, legal, social or organisational issues.

6.2 Additional question(s)

As described in ***Sections 1 and 2***, the aim of the current project is to demonstrate the available evidence on the clinical effectiveness and cost-effectiveness for a broader population than that defined in the Swiss reimbursement text. The following additional questions are contributing to that aim:

1. What is the clinical effectiveness and cost-effectiveness evidence for multigene-expression tests when applied (i.e. for populations and clinical decisions) as described in the reimbursement texts in Switzerland?
2. If no such evidence is available, which application of multigene expression tests for which there is evidence available best matches the Swiss setting?

7. Methodology

The general methodology for the HTA will consist of one systematic review for the clinical evaluation (***Section 7.1***), one systematic review for the economic evaluation (***Section 7.2***), and non-systematic reviews for the evaluation of the ethical, legal, social and organisational domains of the HTA (***Section 7.3***). The proposed methodology for the health economic and budget-impact modelling is outlined in ***Section 7.2.6*** and ***Section 7.2.7***.

7.1 Clinical evaluation

A systematic review is a method to identify, appraise and synthesise all the empirical evidence that meets pre-specified eligibility criteria to answer a specific research question.[44] The systematic review methodology described in this HTA protocol is developed in line with the Cochrane Handbook for Systematic Reviews of Interventions (version 6.4)[44] and the reporting of the systematic review will follow the recommendations of the Preferred Items for Systematic Reviews and Meta-Analyses (PRISMA) Statement.[45]

7.1.1 Databases and search strategy

Systematic literature searches will be conducted in 3 databases: PubMed (MEDLINE), Embase.com and the Cochrane Library. The syntax of the search strategy is composed for one medical database, PubMed (MEDLINE), and customised to the other databases. To gain insight in ongoing RCTs with study characteristics in line with the PICO of this HTA, searches will be conducted on the websites of ClinicalTrials.gov (https://clinicaltrials.gov) and the European Union Clinical Trials Register (https://www.clinicaltrialsregister.eu).

The search strategy is developed with an information specialist based on the PICOs reported in ***Chapter 5***. As quality control the search strategies are checked by a second researcher and validated with a set of key articles. Search strings are compiled for the broad population of patients with breast cancer and the interventions Oncotype DX, MammaPrint, EndoPredict and Prosigna. For PICO 1 a stepwise systematic literature search approach will be implemented: (1) a systematic literature search for prospective RCTs; and (2) in case less than one prospective RCT is found for a multigene-expression test for the clinical outcomes, an additional systematic literature search for retrospective/re-analyses of RCTs and comparative non-randomised studies will be conducted. For this HTA topic, prospective RCTs as well as retrospective or re-analyses of RCTs have been conducted. The primary interest is in prospective RCTs, which use the multigene-expression test prospectively to guide the treatment decision on adjuvant chemotherapy. These trials will provide the highest quality of the available evidence. In contrast, in retrospective RCTs multigene-expression tests are performed on stored resected breast tissue samples. Search limits are added for a publication date limit of the last 15 years and to exclude conference abstracts and preprints (i.e. preprints are excluded since these are not peer-reviewed and the results reported in the preprints might deviate from the final published peer-reviewed article). The details of the search strategies are outlined in ***Appendix B***.

Electronic records of the articles retrieved by the searches will be stored by using Endnote reference manager software (Clarivate Analytics, United States of America [USA]). This Endnote file will be uploaded in Rayyan software (Rayyan Systems Inc., USA) for the selection of the articles.[46] Duplicate records will be deleted and this number is registered in the PRISMA flow diagram.

7.1.2 Study selection

Relevant articles will be selected in duplicate by a systematic approach by 2 independent researchers. Firstly, the major topics of the articles will be assessed on relevancy to the objectives by the title and abstract. Articles that seem to contain relevant data for the objectives will be selected for full-text screening. Articles without relevancy to the objectives will be excluded, without documenting the reason for exclusion. If the 2 researchers disagree on the relevance of an article, this will be discussed. If the differences remain after discussion, the article will be assessed in full text. Secondly, the articles will be assessed in full text based on the pre-specified inclusion criteria (***Table 2***), based on elements of the article, study design and PICOs (***Chapter 5***). If relevant additional criteria emerge during the study selection, this table will be complemented in close collaboration with the FOPH. The final list of applied inclusion criteria will be presented in the HTA report, acknowledging any post-protocol additions or modifications. Articles will be included in the systematic review if they fulfil the inclusion criteria; the remaining articles will be excluded and the primary reason for exclusion is listed. Any differences between the 2 researchers will be resolved by discussion, if needed a third researcher is consulted.

Table 2. Inclusion criteria for clinical studies

	PICO 1 – RCTs		PICO 1 – retrospective/re-analyses of RCTs and comparative non-randomised studies		PICO 2 – comparative non-randomised studies
Publication year	last 15 years (≥2009)				
Language	English, French, German, Italian				
Country of study	worldwide		worldwide		Western countries[a]
Study design/ publication type	- prospective RCTs - systematic reviews (only used for a reference check)		- retrospective/re-analyses of RCTs - longitudinal studies with ≥2 study arms and parallel follow-up, in which 1 arm is exposed and 1 arm is not[b] - systematic reviews (only used for a reference check)		- before-and-after studies - systematic reviews (only used for a reference check)
Population	women and men with ER+, HER2-, LN0-3 early breast cancer after surgical resection[c]	women and men with ER+, HER2-, LN0-3 early breast cancer after surgical resection[c] with multigene-expression tests	women and men with ER+, HER2-, LN0-3 early breast cancer after surgical resection[c]	women and men with ER+, HER2-, LN0-3 early breast cancer after surgical resection[c] with multigene-expression tests	women and men with ER+, HER2-, LN0-3 early breast cancer after surgical resection[c]
Intervention	conventional testing[d] and multigene-expression tests (i.e. Oncotype DX, MammaPrint, EndoPredict, Prosigna)	- chemo-endocrine therapy - chemotherapy	conventional testing[d] and multigene-expression tests (i.e. Oncotype DX, MammaPrint, EndoPredict, Prosigna)	- chemo-endocrine therapy - chemotherapy	conventional testing[d] and multigene-expression tests (i.e. Oncotype DX, MammaPrint, EndoPredict, Prosigna)
Comparator	conventional testing[d]	no chemotherapy	conventional testing[d]	no chemotherapy	conventional testing[d]
Outcome[e]	- overall survival - disease-free survival - recurrence-free-survival - HRQoL		- overall survival - disease-free survival - recurrence-free-survival - HRQoL		impact on treatment management
Other	- articles with unique data (article with the largest sample size or most extended follow-up will be included)[f] - follow-up ≥5 years		- articles with unique data (article with the largest sample size or most extended follow-up will be included)[f] - follow-up ≥5 years - sample size ≥100 - adjusting for main confounders (i.e. age, tumour stage, differences in chemotherapy)		sample size ≥100

Abbreviations:
ER+ = oestrogen receptor positive, HER2- = human epidermal growth factor receptor 2 negative, HRQoL = health-related quality of life, LN0-3 = up to 3 positive lymph nodes, RCTs = randomised controlled trials.

Notes:
a = Austria, Australia, Belgium, Bulgaria, Canada, Croatia, Cyprus, Czechia, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Japan, Latvia, Lithuania, Luxembourg, Malta, the Netherlands, New Zealand, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, Switzerland, United Kingdom, United States of America.[47]
b = In case less than 1 RCT is found for a multigene-expression test, a second systematic literature search step for comparative non-randomised studies will be conducted.
c = Studies are eligible for inclusion if they include the population of interest, present a subgroup analysis for the population of interest, or include a mixed population. In case of a mixed population, the study is eligible for inclusion only if the majority of the study population fits the population of interest (i.e. to be determined during the project).
d = With or without clinical prediction tools, such as Adjuvant!Online.
e = Data will be extracted as defined by the authors, including the definition of the outcome.
f = If applicable, unique results from interim studies will be included (e.g. when results on an outcome of interest are reported only in an interim study) and interim studies might be used as additional input on the study methodology.

To provide insight in the details of the selection process, a PRISMA flow diagram with the results of the study selection and a table with the primary reasons for exclusion for each excluded article at full-text review will be composed and included in the HTA report.

Reference lists of relevant systematic reviews to the research question identified during the title and abstract screening will be checked for potentially relevant additional references of primary studies. Narrative reviews will be excluded directly and not be checked for references. In addition, the supplementary search technique backward citation chasing will be applied, i.e. by finding other studies cited within the included articles. All the additionally found primary studies will be assessed based on the pre-specified eligibility criteria.

7.1.3 Data extraction

Relevant data from the included studies found in the medical journal databases will be extracted by one researcher into a standardised data-extraction spreadsheet in Microsoft Excel. The data-extraction spreadsheet will be fully reviewed by a second researcher, differences are resolved by discussion and in case of discrepancy a third researcher is consulted to reach consensus. This spreadsheet will include:

– bibliographic reference
– study characteristics: study design, trial name, study objective, country, setting, study period, length of follow-up, inclusion/exclusion criteria, source of funding
– study population: diagnosis (hormone receptor status, lymph node status, tumour size, tumour staging), sample size, age, sex, menopausal status, clinical risk as defined by conventional testing (including clinical prediction tools if applicable)
– treatment: type, dose, duration
– intervention: type of multigene-expression test, definition risk categories and thresholds
– comparator: type of conventional tests, definition risk categories and thresholds
– outcomes: definitions and results of the outcomes overall survival, disease-free survival, recurrence-free-survival, HRQoL, impact on treatment management
– additional comments: study limitations or issues that will need to be considered for data interpretation.

Details of ongoing RCTs found in ClinicalTrials.gov and the European Union Clinical Trials Register will be extracted in a summary table in Microsoft Word:

– trail registry identification number
– country
– population
– sample size

– intervention

– comparator

– outcomes

– trial status (e.g. recruiting)

– estimated completion date of the trial.

7.1.4 Analysis of study quality

7.1.4.1 Risk of bias of the reported outcomes in the included studies

The included studies will be critically appraised by one researcher using different tools depending on the study design. The critical appraisal will be fully reviewed by a second researcher, differences are resolved by discussion and in case of discrepancy a third researcher is consulted to reach consensus. The quality of RCTs will be assessed with the revised Cochrane Risk of Bias tool for randomised trials (RoB 2).[44,48] The comparative non-randomised studies will be assessed with the Risk of Bias in Non-randomised Studies - of Interventions (ROBINS-I) tool.[49] The risk of bias assessments will be visualised in plots with the web application Robvis.[50]

7.1.4.2 Overall certainty of the evidence

The overall certainty of the evidence on outcome level will be appraised by one researcher using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach.[44,51] The GRADE assessment will be fully reviewed by a second researcher, differences are resolved by discussion and in case of discrepancy a third researcher is consulted to reach consensus. The certainty of a body of evidence is defined as the extent to which one can be confident that the estimated effect of an intervention is close to the true effect. A GRADE assessment of this certainty involves appraisal of 5 domains: (1) risk of bias (i.e. study limitations; as assessed with the RoB 2 tool and ROBINS-I tool), (2) inconsistency (i.e. heterogeneity or variability in the estimates of treatment effect across studies), (3) indirectness of evidence (i.e. the degree of differences between the PICOs of this HTA and the PICOs of the primary studies), (4) imprecision of the effect estimates, and (5) the risk of publication bias. If applicable, 3 domains can upgrade the certainty of evidence of comparative non-randomised studies (i.e. a large effect, a dose-response gradient or plausible residual opposing confounding). Based on the assessments for each domain, the overall evaluation of the certainty of the evidence per outcome can be classified as high, moderate, low or very low. The overall certainty of the evidence will be summarised in a GRADE summary of findings table for each multigene-expression test, together with key information concerning the magnitudes of relative and absolute effects of the intervention and the amount of available evidence.[44,51] GRADEpro GDT software (Evidence Prime Inc., Canada) will be used to construct the summary of findings table for the outcomes overall survival, disease-free survival, recurrence-free-survival, HRQoL, and impact on treatment management.[52]

7.1.5 Data analysis and synthesis

The extracted data of the included studies in the Microsoft Excel spreadsheet will be summarised in study characteristics tables, risk of bias figures, summary tables, and GRADE summary of findings tables. When possible, forest plots will be created to visualise the outcome results. These tables and figures will be fully reviewed by a second researcher, differences are resolved by discussion and in case of discrepancy a third researcher is consulted to reach consensus.

Based on the comparability of the study populations, interventions and comparators and whether or not data is reported for stratified groups, the options for clinically relevant data merging in a larger group or stratification in subgroups will be explored and, if necessary, discussed with a clinical expert. The clinical expert will be blinded for the study results during this process. Subgroups of populations will be considered, for example the “unclear” population or based on sex; the feasibility depends amongst others on the level of detail about the population within the included studies. The details of the applied data merging/stratification will be reported in the methodology section of the HTA report. The 4 multigene-expression tests will be analysed on an individual level per test, not as one class of the 4 multigene-expression tests combined.

Pooled estimates will be calculated by meta-analysis for outcomes, provided the outcome measures can be combined. If applicable, separate pooled estimates will be calculated for the outcome data of RCTs and the comparative non-randomised studies. Considering the expected heterogeneity in the data, a random-effects model (i.e. DerSimonian and Laird) will be used for the analyses.[38] Heterogeneity will be assessed graphically with forest plots and statistically using the Chi2 test, the I2 statistic and prediction intervals. The analyses will be conducted with the Comprehensive Meta-Analysis (CMA) software (Bio-stat, USA).[53]

Outcomes for which it is not possible to calculate pooled estimates will be analysed narratively and presented in summary tables and GRADE summary of findings tables. A range of relative effects from the studies or direction of the effect will be presented.

7.2 Economic evaluation

The context and rationale for the clinical evaluation as described in ***Section 7.1*** also holds for the economic evaluation. Therefore, the general methodological approach will be the same for the clinical and the economic evaluation. The details of the systematic review methodology for the economic evaluation are described in the next sections.

7.2.1 Databases and search strategy

The cost-effectiveness systematic literature search follows the principles of the systematic literature search for the clinical evaluation outlined in ***Section 7.1.1***. PubMed (MEDLINE), Embase.com and Cochrane library databases will be searched for peer-reviewed scientific literature. In addition, economic databases (the Tufts Medical Centre Cost-Effectiveness Analysis [CEA] Registry and the National Health Service Economic Evaluation Database [NHS EED]) (last update on the 31st of

March 2015) will be searched. The searches will be built using the PICO reported in ***Section 5***. In PubMed (MEDLINE), Embase.com and Cochrane library, the search terms of the clinical efficacy and effectiveness literature search will be combined with cost-effectiveness search terms. The details of the search strategy are presented in ***Appendix C***.

7.2.2 Study selection

All articles retrieved from PubMed (MEDLINE), Embase.com, Cochrane library, CEA Registry and NHS EED will be reviewed in duplicate by 2 independent researchers in a similar manner to the systematic approach described in ***Section 7.1.2***, including firstly screening title and abstract and subsequently full-text screening. If the 2 researchers disagree on the relevance of an article, this will be discussed. If the differences remain after discussion, the article will be assessed in full text. In the first step, the major topics of the articles will be assessed based on relevancy and articles that seem to contain relevant data for the HTA objectives will be selected for the full-text screening. Subsequently, the articles screened in full text will be assessed for inclusion based on pre-specified inclusion criteria defined in ***Table 3***. Like with the clinical evaluation eligibility criteria, if any relevant additional criteria emerge during the study selection, this table will be complemented in close collaboration with the FOPH, and any post-protocol modifications or additions will be described as such in the final report. The screening of full-text articles will be done in duplicate by 2 independent researchers. Any differences will be resolved by discussion, if needed a third researcher is consulted. The final list of applied inclusion criteria will be presented in the HTA report. Articles will be included in the systematic review if they fulfil the inclusion criteria; the remaining articles will be excluded and the primary reason for exclusion is listed. Any differences between the 2 researchers will be resolved by discussion, if needed a third researcher is consulted.

The process of selection and inclusion and exclusion of articles will be recorded with Rayyan software (Rayyan Systems Inc., USA) and in Endnote. This method will provide transparency regarding all selection steps and assures reproducibility. The selection procedure applied during the full-text screening phase will be reported in a PRISMA flow diagram and primary reasons for exclusion per excluded article are listed in a table, like in the clinical evaluation approach.

Table 3: Inclusion criteria for economic studies

Publication year	last 15 years (≥2009)
Language of publication	English, French, German, Italian
Country of study	worldwide
Study design[a]/ publication type	cost-utility analysis cost-effectiveness analysis cost-minimisation analysis cost-benefit analysis

	cost-consequence analysis costing studies	
Population	women and men with ER+, HER2-, LN0-3 early breast cancer after surgical resection[b]	women and men with ER+, HER2-, LN0-3 early breast cancer after surgical resection[b] with multigene-expression tests
Intervention	conventional testing[c] and multigene-expression tests (i.e. Oncotype DX, MammaPrint, EndoPredict, Prosigna)	- chemo-endocrine therapy - chemotherapy
Comparator	conventional testing[c]	no chemotherapy
Outcome	incremental/ total healthcare costs incremental/ total life years and QALYs ICER budget impact	
Other	articles with unique data (article with the largest sample size or most extended follow-up will be included)[d] follow-up ≥5 years[e]	

Abbreviations:
ER+ = oestrogen receptor positive, HER2- = human epidermal growth factor receptor 2 negative, ICER = incremental cost-effectiveness ratio, LN0-3 = up to 3 positive lymph nodes, QALY = quality adjusted life year

Notes:
a = Studies on resource use measurement and early health technology assessment/early economic evaluation are out of scope and will be excluded.
b = Studies are eligible for inclusion if they include the population of interest, present a subgroup analysis for the population of interest, or include a mixed population. In case of a mixed population, the study is eligible for inclusion only if the majority of the study population fits the population of interest (i.e. to be determined during the project).
c = With or without clinical prediction tools, such as Adjuvant!Online.
d = If applicable, unique results from interim studies will be included (e.g. when results on an outcome of interest are reported only in an interim study) and interim studies might be used as additional input on the study methodology.
e = This relates to the follow-up time point of the clinical study that the economic study is based on.

7.2.3 Data extraction

Relevant data from the included studies found in the medical journal databases will be extracted by one researcher into a standardised data-extraction spreadsheet in Microsoft Excel. The data-extraction spreadsheet will be fully reviewed by a second researcher, differences are resolved by discussion and in case of discrepancy a third researcher is consulted to reach consensus. This spreadsheet will include:

– first author, year

– country

– type of study
– study perspective
– time horizon
– discount rate
– study funding
– study population (sample size, mean age, age range, proportion men/ women, menopausal status)
– intervention (testing and treatment)
– comparator (testing and treatment)
– outcomes: definitions and results of the outcomes; total/ incremental costs and quality-adjusted life years (QALYs)
– model used (yes/no, type of model and health states)
– primary sources for the resource use/cost inputs
– primary sources for the HRQoL inputs.

7.2.4 Analysis of study quality

The identified studies from the systematic literature search for cost-effectiveness will be subjected to a critical appraisal using the Consolidated Health Economic Evaluation Reporting Standards (CHEERs)[54] checklist and/or the Consensus Health Economic Criteria (CHEC) checklist[55] as recommended by the current guidelines.[56] The CHEERS and CHEC are 28-item and 19-item checklists, respectively, with clear questions about the economic evaluation that gives insight into the general quality of the study.

7.2.5 Data analysis and synthesis

Data synthesis will be done using descriptive comparisons of the study question, methods and results. Summary tables will be included which will present key information described in the data extraction (see ***Section 7.2.3***). The incremental cost-effectiveness ratios (ICERs) will be presented and the reliability (internal validity) and relevance (generalisability) of the estimates will be explored applying the appraisal tools described in ***Section 7.2.4***. The analytical approaches used in the studies will be compared and their robustness is discussed.

7.2.6 Economic model protocol

The development of a health economic model to synthesise the collected data is conditional on the success of finding data relevant to the Swiss setting. If the available data that is most relevant to the Swiss clinical setting (as described in ***Sections 7.1.2 and 7.2.2***) does not allow for a meaningful

economic evaluation of the multigene-expression tests in the Swiss clinical setting, no model will be developed. Whether this criterion is met will be decided in close collaboration with the FOPH. Below is a description of the model characteristics in case it is decided an economic model will be developed.

7.2.6.1 Target population

Women and men with ER+, HER2- and up to 3 positive lymph nodes early breast cancer after surgical resection for whom it is unclear based on conventional testing, whether to prescribe adjuvant chemotherapy.

7.2.6.2 Setting and location

The analysis will be performed for the Swiss healthcare setting. This means that where possible and relevant input parameters will be based on data from Switzerland (e.g. Swiss lifetables for background mortality and Swiss sources for healthcare costs).

7.2.6.3 Study perspective

The analysis will be performed from a healthcare payers' perspective. Costs of healthcare services covered by the Swiss mandatory health insurance will be analysed, irrespective of the actual payer (mandatory health insurer, other social insurer, government (federal government, cantons, communities) out-of-pocket). The analysis will not include indirect costs due to informal care or productivity losses and additional non-medical costs for patients, such as travel costs.

7.2.6.4 Interventions

The intervention of interest are 4 multigene-expression tests. These are:

– Oncotype DX

– MammaPrint

– EndoPredict

– Prosigna

The decision to provide adjuvant chemotherapy will be based on the test outcome using the manufacturer's prescribed outcome categories or cut-offs.

7.2.6.5 Comparators

The comparison for the intervention is conventional testing only, without the use of any clinical prediction tools (i.e. the current standard of care in Switzerland). This includes of the following tests and examinations:

– tumour stage

– nodal stage

– tumour grade

– resection completeness
– lymphatic vessel and vascular invasion
– Ki-67
– oestrogen and progesterone receptor assessment
– HER2 score
– histological subtypes other than ductal carcinoma
– family history
– patient fitness (age, co-morbidities)
– menopausal status
– imaging results

7.2.6.6 Time horizon

The preferred time horizon of the base-case analysis is lifetime. The feasibility of implementing a lifetime horizon will depend on the availability of data. Shorter time horizons will be considered in scenario analyses, if relevant.

7.2.6.7 Discount rate

In the base-case analysis, costs and effects will be discounted at 3.0%. In scenario analyses, the impact of not discounting or using a discount rate of 5.0% will be explored.

7.2.6.8 Health outcomes

Health outcomes will be reported in life years and QALYs.

7.2.6.9 Currency, price data, and conversion

Costs from the Swiss Federal Statistical Office will be reported in Swiss francs (CHF) adjusted for inflation to 2022 price levels using inflation rates, which will be accessed from the Organization for Economic Co-operation and Development website (https://data.oecd.org). Should price index data for 2023 be available at the time of analysis, 2023 prices levels will be used.

7.2.6.10 Model structure

The model structure will be chosen based on the data available to inform model parameters. In general, cost-effectiveness analyses of multigene-expression tests make use of a state transition model consisting of at least 3 health states: ‘Disease-free’, ‘Any recurrence’, and ‘Dead’.[57] In some analyses additional states are used, for example to distinguish between local and distant recurrence.[57]

Depending on the expected complexity of the model, it will either be developed in Microsoft Excel or programmed in statistical programming language R[58] based on the framework developed by the Decision Analysis in R for Technologies in Health workgroup.[59] This decision will be made as the

start of the modelling phase in consultation with the FOPH, taking into account the efficiency of model development and end user functionality.

7.2.6.11 Input parameters

The model input parameters on clinical effects (i.e. recurrence rates, mortality), and HRQoL will be informed by the results of the clinical and economic evaluations as described in ***Sections 7.1*** and above in this section. Background mortality will be based on Swiss lifetables.[60]

All costs within the healthcare sector will be considered, such as the costs of the tests, drug costs, disease monitoring costs and (if modelled) the costs of adverse events. Where possible, Swiss resource use will be used. If not available, international data on resource use will be used instead, multiplied with Swiss unit costs as supplied by the FOPH. If resource use data is not available, international cost estimates will be used.

If no Swiss-specific data on costs, resource use and utilities are identified in the results of the economic literature search, additional pragmatic searches will be performed. Clinical expert opinion will be used whenever data is unavailable from the literature.

7.2.6.12 Analytical methods

7.2.6.12.1 Base case, subgroup, and scenario analyses

The base-case analysis will consider the entire target population and will be conducted as described in the previous sections.

In principle, no subgroup analyses will be conducted, unless a subgroup of critical relevance is identified in the clinical or economic evaluation. This will be decided in close collaboration with the FOPH at the start of the model development.

Structural uncertainty will be explored in several scenario analyses, using alternative assumptions and sources compared to the base case.

7.2.6.12.2 One-way sensitivity analyses

Parameter uncertainty is first tested using one-way sensitivity analyses; model parameters are systematically and independently varied over a plausible range (e.g. using the 95% confidence interval or a 20% increase/decrease of the parameter value used in the base-case). The ICER is recorded at the upper and lower limits to produce tornado diagrams.

7.2.6.12.3 Probabilistic sensitivity analysis

Joint parameter uncertainty is explored through probabilistic sensitivity analysis where all parameters, to which probability distributions are assigned, are varied jointly. Monte Carlo simulations will be performed, and the results are recorded. Results will be plotted on the cost-effectiveness plane. From these results, a cost-effectiveness acceptability curve will be estimated.

7.2.7 Budget impact analysis

In addition to the cost-effectiveness model, a budget impact model will be developed to estimate the projected population-level overall costs of the use of multigene-expression tests to guide adjuvant chemotherapy in early breast cancer. The budget impact model will be built as an extension to the cost-effectiveness model, described above. Hence, the core model characteristics for the budget impact model will be dependent on the cost-effectiveness model. The time horizon will be restricted to 5 years for the period 2025-2029. For the budget impact model, data is required about the current use of multigene-expression tests in Switzerland. This data will be supplied by the FOPH if available. If such data is not available, assumptions will be made based on expert opinion. Uncertainty in these assumptions can be explored through scenario analysis.

7.3 Evaluation of the ethical, legal, social and organisational domains (ELSO)

Studies retrieved in the systematic literature searches for the clinical and economic evaluation will be screened for ELSO and indirect safety issues. In addition, targeted non-systematic searches will be conducted in PubMed (MEDLINE) and on relevant websites, such as the websites of the European Society of Breast Cancer Specialists (eusoma.org), the Schweizerische Gesellschaft für Senologie (senologie.ch), the Union for International Cancer Control (uicc.org), and other websites that may be identified. The selection procedure applied during the full-text screening phase will be reported in a PRISMA flow diagram and key findings will be presented narratively.

8. Summary and Outlook

8.1 Summary

In Switzerland, multigene-expression tests are temporarily covered by the mandatory health insurance for patients with early ER+, HER2- breast cancer, with up to 3 affected lymph nodes, for whom the decision to prescribe adjuvant chemotherapy is unclear. The question whether coverage of the multigene-expression tests (i.e. Oncotype DX, MammaPrint, EndoPredict and Prosigna) should be extended, is to be re-evaluated. This HTA is to be conducted to inform policy makers on that decision.

In this HTA protocol, the HTA questions are formulated, the patient, intervention, comparator, and outcomes are defined, and the methodology to conduct the evaluation is described. The methodology consists of 2 systematic literature searches: one for the clinical evaluation and one for the economic evaluation. Following these searches, relevant studies will be selected, the methodological quality of included studies will be critically appraised and data is extracted. The overall certainty of the evidence on outcome level will be assessed with GRADE. If the retrieved data permits, an economic model will be developed to perform a cost-effectiveness analysis. Lastly, a budget impact analysis and an evaluation of ethical, legal, social and organisational domains will be conducted.

A number of challenges during the HTA phase is foreseen. Given the large amount of evidence being published over the many years of research in this field, it is expected that findings will display a substantial degree of heterogeneity in terms of patient population, testing strategy and use of test results, comparator, outcome, study design, length of follow-up and others. This heterogeneity will be identified, presented, and discussed in the subsequent HTA report.

Furthermore, there is a substantial degree of discordance between tests. The discordance is related to the differences between the tests in terms of gene selection, stage of development, patient population, adjuvant diagnostic tools and interpretation bias. The consulted clinicians pointed out that the choice of tests is often driven by circumstantial factors such as experience of the clinician, hospital policy and financial arguments. For these reasons the issue of discordance between test modalities is not addressed in this HTA (i.e. no test-to-test comparisons). This HTA will provide the best available evidence of effectiveness and cost-effectiveness per test.

8.2 Outlook

The HTA protocol is followed by an HTA report. The objective of the HTA report is to generate a focused assessment of various aspects of the health technology in question. The applied analytic methods, their execution and the results are described. The analytical process is comparative, systematic and transparent. The external review group that was consulted during the protocol phase is consulted again during the HTA phase. Subsequently the HTA draft report is presented to the stakeholders for consultation. Communication with the reviewers and the stakeholders is coordinated by the FOPH.

9. References

1. Burstein HJ, Curigliano G, Thürlimann B, et al. Customizing local and systemic therapies for women with early breast cancer: the St. Gallen International Consensus Guidelines for treatment of early breast cancer 2021. *Annals of Oncology* 2021; 32: 1216–1235.

2. Curigliano G, Burstein HJ, Gnant M, et al. Understanding breast cancer complexity to improve patient outcomes: The St Gallen International Consensus Conference for the Primary Therapy of Individuals with Early Breast Cancer 2023. *Annals of Oncology* 2023; 34: 970–986.

3. Loibl S, André F, Bachelot T, et al. Early breast cancer: ESMO Clinical Practice Guideline for diagnosis, treatment and follow-up. *Annals of Oncology* 2024; 35: 159–182.

4. Krebsliga. Krebs in der Schweiz: wichtige Zahlen. https://www.krebsliga.ch/ueber-krebs/zahlen-fakten/ (accessed 07 June 2023).

5. Wanner M, Matthes KL, Karavasiloglou N, et al. 37-year incidence and mortality time trends of common cancer types by sex, age, and stage in the canton of Zurich. *Swiss Med Wkly* 2020; 150: w20388.

6. National Agency for Cancer Registration. Le Cancer en Suisse—Rapport 2021., https://www.onec.ch/fr/population/le-cancer-en-suisse-rapport-2021/ (accessed 16 June 2023).

7. Feng Y, Spezia M, Huang S, et al. Breast cancer development and progression: Risk factors, cancer stem cells, signaling pathways, genomics, and molecular pathogenesis. *Genes & Diseases* 2018; 5: 77–106.

8. Waks AG, Winer EP. Breast Cancer Treatment: A Review. *JAMA* 2019; 321: 288.

9. Howlader N, Altekruse SF, Li CI, et al. US Incidence of Breast Cancer Subtypes Defined by Joint Hormone Receptor and HER2 Status. *JNCI: Journal of the National Cancer Institute*; 106. Epub ahead of print May 2014. DOI: 10.1093/jnci/dju055.

10. Breast cancer. *World Health Organization*, https://www.who.int/news-room/fact-sheets/detail/breast-cancer (accessed 11 September 2023).

11. McDonald ES, Clark AS, Tchou J, et al. Clinical Diagnosis and Management of Breast Cancer. *J Nucl Med* 2016; 57: 9S-16S.

12. Breast Cancer Stages. *National Breast Cancer Foundation*, https://www.nationalbreastcancer.org/about-breast-cancer/breast-cancer-staging (accessed 27 September 2023).

13. Cardoso F, Kyriakides S, Ohno S, et al. Early breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Annals of Oncology* 2019; 30: 1194–1220.

14. Gerdes J, Schwab U, Lemke H, et al. Production of a mouse monoclonal antibody reactive with a human nuclear antigen associated with cell proliferation. *Int J Cancer* 1983; 31: 13–20.

15. Salvo EM, Ramirez AO, Cueto J, et al. Risk of recurrence among patients with HR-positive, HER2-negative, early breast cancer receiving adjuvant endocrine therapy: A systematic review and meta-analysis. *The Breast* 2021; 57: 5–17.

16. Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Comparisons between different polychemotherapy regimens for early breast cancer: meta-analyses of long-term outcome among 100 000 women in 123 randomised trials. *The Lancet* 2012; 379: 432–444.

17. Anampa J, Makower D, Sparano JA. Progress in adjuvant chemotherapy for breast cancer: an overview. *BMC Med* 2015; 13: 195.

18. Giorgi Rossi P, Lebeau A, Canelo-Aybar C, et al. Recommendations from the European Commission Initiative on Breast Cancer for multigene testing to guide the use of adjuvant chemotherapy in patients with early breast cancer, hormone receptor positive, HER-2 negative. *Br J Cancer* 2021; 124: 1503–1512.

19. Goldhirsch A, Ingle JN, Gelber RD, et al. Thresholds for therapies: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2009. *Annals of Oncology* 2009; 20: 1319–1329.

20. Paik S, Shak S, Tang G, et al. A Multigene Assay to Predict Recurrence of Tamoxifen-Treated, Node-Negative Breast Cancer. *N Engl J Med* 2004; 351: 2817–2826.

21. Paik S, Tang G, Shak S, et al. Gene Expression and Benefit of Chemotherapy in Women With Node-Negative, Estrogen Receptor–Positive Breast Cancer. *JCO* 2006; 24: 3726–3734.

22. Dowsett M, Cuzick J, Wale C, et al. Prediction of Risk of Distant Recurrence Using the 21-Gene Recurrence Score in Node-Negative and Node-Positive Postmenopausal Patients With Breast Cancer Treated With Anastrozole or Tamoxifen: A TransATAC Study. *JCO* 2010; 28: 1829–1834.

23. Albain KS, Barlow WE, Shak S, et al. Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal women with node-positive, oestrogen-receptor-positive breast cancer on chemotherapy: a retrospective analysis of a randomised trial. *The Lancet Oncology* 2010; 11: 55–65.

24. Sparano JA, Gray RJ, Makower DF, et al. Adjuvant Chemotherapy Guided by a 21-Gene Expression Assay in Breast Cancer. *N Engl J Med* 2018; 379: 111–121.

25. Sparano JA, Gray RJ, Makower DF, et al. Clinical Outcomes in Early Breast Cancer With a High 21-Gene Recurrence Score of 26 to 100 Assigned to Adjuvant Chemotherapy Plus Endocrine Therapy: A Secondary Analysis of the TAILORx Randomized Clinical Trial. *JAMA Oncol* 2020; 6: 367.

26. Haan JC, Bhaskaran R, Ellappalayam A, et al. MammaPrint and BluePrint comprehensively capture the cancer hallmarks in early-stage breast cancer patients. *Genes Chromosomes Cancer* 2022; 61: 148–160.

27. Buyse M, Loi S, van't Veer L, et al. Validation and Clinical Utility of a 70-Gene Prognostic Signature for Women With Node-Negative Breast Cancer. *JNCI: Journal of the National Cancer Institute* 2006; 98: 1183–1192.

28. On behalf of the TRANSBIG consortium, Mook S, Schmidt MK, et al. The 70-gene prognosis-signature predicts disease outcome in breast cancer patients with 1–3 positive lymph nodes in an independent validation study. *Breast Cancer Res Treat* 2009; 116: 295–302.

29. van de Vijver MJ, He YD, van 't Veer LJ, et al. A Gene-Expression Signature as a Predictor of Survival in Breast Cancer. *N Engl J Med* 2002; 347: 1999–2009.

30. van 't Veer LJ, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002; 415: 530–536.

31. Sapino A, Roepman P, Linn SC, et al. MammaPrint Molecular Diagnostics on Formalin-Fixed, Paraffin-Embedded Tissue. *The Journal of Molecular Diagnostics* 2014; 16: 190–197.

32. Drukker CA, Schmidt MK, Rutgers EJT, et al. Mammographic screening detects low-risk tumor biology breast cancers. *Breast Cancer Res Treat* 2014; 144: 103–111.

33. Győrffy B, Hatzis C, Sanft T, et al. Multigene prognostic tests in breast cancer: past, present, future. *Breast Cancer Res* 2015; 17: 11.

34. Cardoso F, van't Veer LJ, Bogaerts J, et al. 70-Gene Signature as an Aid to Treatment Decisions in Early-Stage Breast Cancer. *N Engl J Med* 2016; 375: 717–729.

35. Soliman H, Shah V, Srkalovic G, et al. MammaPrint guides treatment decisions in breast Cancer: results of the IMPACt trial. *BMC Cancer* 2020; 20: 81.

36. Filipits M, Rudas M, Jakesz R, et al. A New Molecular Predictor of Distant Recurrence in ER-Positive, HER2-Negative Breast Cancer Adds Independent Information to Conventional Clinical Risk Factors. *Clinical Cancer Research* 2011; 17: 6012–6020.

37. Filipits M, Dubsky P, Rudas M, et al. Prediction of Distant Recurrence Using EndoPredict Among Women with ER+, HER2− Node-Positive and Node-Negative Breast Cancer Treated with Endocrine Therapy Only. *Clinical Cancer Research* 2019; 25: 3865–3872.

38. Myriad EndoPredict® Technical Specifications.

39. Buus R, Sestak I, Kronenwett R, et al. Comparison of EndoPredict and EPclin With Oncotype DX Recurrence Score for Prediction of Risk of Distant Recurrence After Endocrine Therapy. *JNCI J Natl Cancer Inst* 2016; 108: djw149.

40. Parker JS, Mullins M, Cheang MCU, et al. Supervised Risk Predictor of Breast Cancer Based on Intrinsic Subtypes. *JCO* 2009; 27: 1160–1167.

41. Dowsett M, Sestak I, Lopez-Knowles E, et al. Comparison of PAM50 Risk of Recurrence Score With Onco *type* DX and IHC4 for Predicting Risk of Distant Recurrence After Endocrine Therapy. *JCO* 2013; 31: 2783–2790.

42. Filipits M, Nielsen TO, Rudas M, et al. The PAM50 Risk-of-Recurrence Score Predicts Risk for Late Distant Recurrence after Endocrine Therapy in Postmenopausal Women with Endocrine-Responsive Early Breast Cancer. *Clinical Cancer Research* 2014; 20: 1298–1305.

43. Lænkholm A-V, Jensen M-B, Eriksen JO, et al. PAM50 Risk of Recurrence Score Predicts 10-Year Distant Recurrence in a Comprehensive Danish Cohort of Postmenopausal Women Allocated to 5 Years of Endocrine Therapy for Hormone Receptor–Positive Early Breast Cancer. *JCO* 2018; 36: 735–740.

44. Higgins J, Thomas J, Chandler J, et al. Cochrane Handbook for Systematic Reviews of Interventions version 6.3 (updated February 2022), www.training.cochrane.org/handbook (2022).

45. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021; n71.

46. Ouzzani M, Hammady H, Fedorowicz Z, et al. Rayyan—a web and mobile app for systematic reviews. *Syst Rev* 2016; 5: 210.

47. United Nations Department of Economic and Social Affairs. *World Economic Situation and Prospects 2023*. United Nations.

48. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019; l4898.

49. Sterne JA, Hernán MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ* 2016; i4919.

50. Risk of bias tools - robvis (visualization tool), https://www.riskofbias.info/welcome/robvis-visualization-tool.

51. Schünemann H, Brożek J, Guyatt G, Oxman A. *Introduction to GRADE Handbook. Handbook for grading the quality of evidence and the strength of recommendations using the GRADE approach*. 2013.

52. GRADEpro, https://www.gradepro.org/ (accessed 17 July 2023).

53. Borenstein M, Hedges L, Higgins J, et al. Comprehensive Meta-Analysis Version 4.

54. Husereau D, Drummond M, Augustovski F, et al. Consolidated Health Economic Evaluation Reporting Standards 2022 (CHEERS 2022) Statement: Updated Reporting Guidance for Health Economic Evaluations. *Value in Health* 2022; 25: 3–9.

55. Evers S, Goossens M, de Vet H, et al. Criteria list for assessment of methodological quality of economic evaluations: Consensus on Health Economic Criteria. *Int J Technol Assess Health Care* 2005; 21: 240–245.

56. Aluko P, Graybill E, Craig D, et al. Chapter 20: Economic evidence. In: *In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.3 (updated February 2022)*. Cochrane, www.training.cochrane.org/handbook (2022).

57. NICE. *Tumour profiling tests to guide adjuvant chemotherapy decisions in people with breast cancer (update of DG10)*. 18 October 2017.

58. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, https://www.R-project.org/ (2021).

59. Alarid-Escudero F, Krijkamp E, Pechlivanoglou P, et al. A need for change! A coding framework for improving transparency in decision modeling. *Pharmacoeconomics* 2019; 37: 1329–1339.

60. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Human Mortality Database, https://www.mortality.org/ (accessed 23 December 2021).

10. Appendices

A. Summary of the clinical experts advisory panel meeting

Minutes by HTA assessment team

Meeting with advisory clinical expert group for the planning of an HTA report on:

"Use of multigene-expression tests to guide decision making on adjuvant chemotherapy in patients with breast cancer"

Thursday, 13th of July 2023

Online via MS-Teams

Attendees	*It was agreed with the participants that the results of the meeting would only be shared anonymously. The identities of the participants are known to the FOPH and contractor.*

Questions about standard of care (SoC)

To our knowledge the typical standard of care (SoC) testing and examinations administered in Swiss clinical practice to patients with early breast cancer after surgery to guide clinicians on adjuvant chemotherapy is based on the risk of tumour recurrence defined by clinical, histological and immunohistochemical parameters.

Histological/ immuno-histochemical parameters	Clinical parameters
Tumour stage	Family history
Nodal stage (N0-3)	Patient fitness (age, co-morbidities)
Tumour grade (G1-3)	Imaging results
Resection completeness (R0-2)	Patient preferences
Lymphatic vessel/ vein invasion (L0-1, V0-1)	
Ki-67	
Hormone receptor assessment (ER%, PgR%)	
HER2 score	

1.1 Are parameters related to standard of care testing and examinations to guide adjuvant chemotherapy missing?

The experts agreed that histological subtypes other than ductal carcinoma (such as lobular or other special subtypes) should be added to the list of histological/ immune-histochemical parameters. In addition, the term 'vein invasion' should be replaced by 'vascular invasion'. Additionally, the tumour-infiltrating lymphocytes (TILs) were considered relevant by some experts. TILs, however, were not considered as part of standard of care, and some experts would be cautious with using this parameter.

The experts agreed that in the clinical parameters list, the term menopausal status should be added.

1.2 What results (i.e. which cut-off values) or combination of results on the parameters contribute to a positive or negative advice regarding adjuvant chemotherapy?

The clinical experts responded that it would not be possible to give a clear answer, because a positive or negative advice would derive from a combination of all those parameters.

In general, the experts agreed that scores on the higher end of the scale of each parameter could categorise the patients as high risk and respectively scores on the lower end would categorise them as low risk. Some examples given by the experts about cases that adjuvant chemotherapy would be given include grade 3 tumours and tumours with high Ki-67 score (e.g. > 30%) whereby the definition of high is still a matter of debate (Ki-67 is, however, agreed to not be reliable by itself).

1.3 When are the standard of care test results considered insufficient or uncertain to recommend adjuvant chemotherapy?

The experts agreed that, individually, all parameters have limitations, therefore a clear answer cannot be given. In general, results in the middle of the range of each parameter are considered to be ambiguous with regard to guiding adjuvant chemotherapy (e.g. tumour size smaller than 1-2 cm, grade 2 tumours).

2. *Are the tests and examinations and the interpretation of their results the same throughout Switzerland?*

The experts agreed with the statement. The experts explained that in certified breast cancer centres these tests and examinations (with the focus on biomarkers), and how they are interpreted are part of the standard of care. Furthermore, these tests are or should be subject to regular quality assurance.

3. *Are clinical prediction tools such as Adjuvant!, the Nottingham Prognostic Index or Predict used to guide adjuvant chemotherapy in Switzerland?*

The experts agreed that the clinical prediction tools are more frequently used to assist in communicating prognosis and potential chemotherapy benefit to the patient, and not to make a decision about whether to prescribe adjuvant chemotherapy. Therefore, most physicians do not use the tests in order to make a decision about adjuvant chemotherapy.

It was noted by one of the experts that the tools might be more frequently used by physicians who are less experienced.

Of the tools mentioned above, the experts agreed that many physicians liked “Adjuvant!” Online in the first place but that the use was discontinued. In contrast, the tool “Predict” is more frequently used, even though it was not regarded as accurate. And as explained in the beginning the purpose of using them is rather to illustrate prognosis and thus communicate prognosis to patients.

Additional question. What percentage of the Swiss practitioners applies those tools?

The experts expressed that this is not known, but the general consensus is that Predict is used quite frequently.

Questions about multigene-expression-tests

To our knowledge, multigene-expression tests are not used in all patients with ER+, HER2-, LN0-3 early breast cancer. General condition of the patient (age, comorbidities) and patient preferences are relevant considerations.

Furthermore, for patients with an assessed **low** risk of recurrence are recommended not to take adjuvant chemotherapy (endocrine therapy only), while patients with a **high** risk of recurrence are recommended to take adjuvant chemotherapy.

For those patients not considered low or high risk, a multigene-expression test can be helpful.

4.1 – A Do you agree with the statements in the blue text box?

The experts agreed that the statements were reasonable. The experts explained that there are indeed patients for whom the risk assessment based on conventional testing only is sufficient to make a decision on adjuvant chemotherapy, and thus that multigene-expression tests are not needed to make that decision.

4.1 – B Do you agree with the definition of low and high risk of recurrence as shown below?

Low risk	High risk
pT1	pT2 or pT3
Grade 1 or 2	Grade 3
ER and PgR highly positive	Positive nodes
Low Ki67	High Ki67
Invasive lobular carcinoma	Patient age <35 years
Elderly patient or patient with comorbidities	

The experts did not entirely agree with the contents of the table, and provided the following remarks:

The experts remarked to better replace the description 'low risk' and 'high risk' by 'leaning towards no chemotherapy' and 'leaning towards chemotherapy'. Invasive lobular carcinoma and elderly patients or patients with comorbidities per se cannot be regarded as low risk factors. If the terminology 'classical lobular' were used, which reflects 80% of the invasive lobular carcinomas, then it could be considered as leaning towards no chemotherapy (as these tumours have a low sensitivity to chemotherapy). Additionally, the size of the tumour has very little prognostic ability, while grade 2 tumours should be moved from low risk to high risk. A low Ki67 has important prognostic ability. Furthermore, positive nodes can apply to a lot of different situations with different meanings. Especially pN2 patients (>3 positive lymph nodes) are not candidates for multigene-expression tests.

4.2 Do you think that this statement (i.e. multigene-expression tests are used in patients that are not classified low or high risk based on conventional testing) accurately reflects Swiss clinical practice?

The experts agreed with the statement. However, one expert added that some centres perform EndoPredict on all patients with ER+, HER2-, LN0-3 early breast cancer.

4.3 How do you interpret the text of the current reimbursement condition «konventionelle Befunde allein erlauben keine eindeutige Entscheidung…»*? Do you think this captures the same patients that fall outside of the previously defined low and high risks groups?*

The experts agreed that the way this statement was formulated is just another way of formulating the grey zone. Even though there could be some variance, the experts agree it is okay to leave some leeway.

Therefore, the experts agreed that this statement can be left as it is, as there is no good reason to widen it without clearly defining the unclear group.

5. *Do clinicians typically use one type of multigene-expression test, or do they use different tests for different patient groups?*

   The experts agreed that the type of test that is used usually depends on the institution, as it is more related to the facilities than the individual patient characteristics (i.e. there is no evidence available that test x performs better in patient y than test z).

Additional question: Are the tests interchangeable?

The experts considered the tests as non-interchangeable. One expert explained that Oncotype has the best evidence; the test was shown to be predictive for the benefit of chemotherapy (the other tests have evidence on prognostic ability only). Furthermore, the Oncotype gene expression test showed in a randomized phase 3 trial that, at least in postmenopausal patients, the intermediate risk group identified by the Oncotype test did not benefit from chemotherapy. Some experts added that there were instances where two tests were performed, when two separate physicians prescribed them, which can lead to contradicting results.

6. *Are the multigene-expression tests used only on top of SoC testing, or do they replace part of SoC testing?*

   All experts indicated that multigene-expression tests do not replace part of SoC testing but are added on top. The decision to use a multigene test is based on the pathology report. The experts remark that oncologists should be aware of mean and median Ki67 values of the pathology institutions they work with when interpreting results, as this parameter can vary between the institutions.

7. *What is the weight of the multigene-expression tests findings in decision making on whether to recommend adjuvant chemotherapy? Are the results considered as additional confirmative information or as more directive information?*

   The experts agreed that the results of the test are considered in a directive way. They explained that in situations where there is uncertainty about the benefit of adjuvant chemotherapy, and the tests highly recommend adjuvant chemotherapy, then the physician should take that direction. It would be pointless to perform a test if the direction of the results would not be followed.

Additional question 1: Are multigene-expression tests performed after all conventional testing is already done?

The experts acknowledged that this is indeed the case.

Additional question 2: Do you always follow the cut-off point of the manufacturer?

The experts agreed that there would be some grey zones. One expert gave the example that with respect to Oncotype a controversy exists between pre- and post-menopausal women when choosing which cut-off point to use.

Questions about the planned HTA report

8. *For the planned HTA report does the advisory clinical expert group agree on the following population?*

   Patients with oestrogen receptor positive (ER+), human epidermal growth factor receptor 2 negative (HER2-) and up to 3 positive nodes (LN0-3) early breast cancer after surgical resection for whom it is unclear based on conventional testing, whether to prescribe adjuvant chemotherapy.

   The experts agreed with the described population, however one expert explained that for patients with negative nodes and very small tumours (< 1cm), the effect of chemotherapy is very small. In these patients with very small tumours the tests do not have the same value.

Additionally, one expert advised that, if possible, a subgroup analysis should be conducted in these patients.

9. *In the planned HTA report, multigene-expression tests will be compared to using no multigene-expression tests. Does the advisory clinical expert group also consider a comparison among the different multigene-expression tests as relevant?*

   The experts agreed that a comparison among multigene-expression tests would not be possible, and they advised not to attempt it, as only the results of prospective trials could give answers and these are lacking for the majority of tests (it was mentioned that at least one such trial is ongoing, so more evidence can be expected to be available in the future). One expert emphasised that it is important to report the level of evidence per test (as is already included in the AGO and NCCN guidelines), and it should be acknowledged that there are differences in populations across studies.

10. *For the planned HTA report would the advisory clinical expert group recommend to agree on a minimal important clinical difference (MCID) for the outcome recurrence rate (e.g. 5% absolute difference), or is this considered impossible or not constructive?*

    The experts agreed that they would not consider it as constructive (populations across studies/ tests vary which influences the MCID). Additionally, one expert pointed out that Oncotype is the only multigene-expression test that provides predictive results. Finally, one expert explained that there is a variety of different recurrence types that can occur, but distant metastases seem the most relevant recurrence type in this context.

11. *Given the planned HTA, is an important clinical question missing in order to understand the clinical evidence?*

    The experts agreed that everything was covered.

    The following suggestions were given by some of the experts:

- Presenting data on how the usage trends of multigene-expression tests in Switzerland changed over the years, and which patient groups the multigene-expression tests are used on.
- Reassessing the value of multigene-expression test (in terms of QALYs and cost-effectiveness), since now the tests are being used in the real world and not in a trial setting.

Additional question: Would multigene tests prevent from second opinions or is that such a common practice that tests would not change anything on that aspect?

The experts agreed that second opinion consultations would not really be affected since they are different things: the result of a multigene-expression test is mostly not the reason for a second opinion. Moreover, the experts explained that second opinions are applied mainly to convince the patients and not the practitioners. They strongly agreed that second opinions are too cheap compared to technical tests/investigations.

Additional question: In case of Oncotype, do you need to send samples to the US in order to get results? *(Please answer this question it was not part of the meeting)*

It was mentioned that there is at least one central lab in Germany. However, it is unclear whether pathologists from outside Germany are allowed to send samples there. One expert indicated that for the physician ordering the test it is of minor importance whether the test needs to be sent to a country in Europe or the US. Instead, timely service and reliability are important. Another expert indicated that sending the samples to the US is fast.

B. Search strategy for clinical evaluation systematic literature search

Table 4: PubMed (MEDLINE)

Population	"Breast Neoplasms"[Mesh] OR breast neoplasm*[tiab] OR breast tumor*[tiab] OR breast tumour*[tiab] OR breast cancer[tiab] OR breast malignanc*[tiab] OR breast carcinoma*[tiab] OR breast adenocarcinoma*[tiab] OR breast sarcoma*[tiab] OR mammary neoplasm*[tiab] OR mammary tumor*[tiab] OR mammary tumour*[tiab] OR mammary cancer[tiab] OR mammary malignanc*[tiab] OR mammary carcinoma*[tiab] OR mammary adenocarcinoma*[tiab] OR mammary sarcoma*[tiab]
Intervention	endopredict*[tiab] OR EPclin[tiab] OR 12-gene[tiab] OR gene-12[tiab] OR oncotype DX*[tiab] OR oncotypeDX*[tiab] OR 21-gene[tiab] OR gene-21[tiab] OR mammaprint*[tiab] OR mamma-print*[tiab] OR 70-gene[tiab] OR gene-70[tiab] OR prosigna*[tiab] OR microarray 50[tiab] OR PAM50[tiab] OR PAM 50[tiab] OR 50-gene[tiab] OR gene-50[tiab]
Comparator	No search string
Outcomes	No search string
Limits	*Publication period* last 15 years (≥2009)
	No conference abstracts and preprints NOT (congress[pt] OR preprint[pt])

Table 5: Embase.com

Population	'breast tumor'/exp OR 'breast neoplasm*':ti,ab OR 'breast tumor*':ti,ab OR 'breast tumour*':ti,ab OR 'breast cancer':ti,ab OR 'breast malignanc*':ti,ab OR 'breast carcinoma*':ti,ab OR 'breast adenocarcinoma*':ti,ab OR 'breast sarcoma*':ti,ab OR 'mammary neoplasm*':ti,ab OR 'mammary tumor*':ti,ab OR 'mammary tumour*':ti,ab OR 'mammary cancer':ti,ab OR 'mammary malignanc*':ti,ab OR 'mammary carcinoma*':ti,ab OR 'mammary adenocarcinoma*':ti,ab OR 'mammary sarcoma*':ti,ab
Intervention	'breast cancer prognostic test kit'/exp OR endopredict*:ti,ab OR EPclin:ti,ab OR 12-gene:ti,ab OR gene-12:ti,ab OR 'oncotype DX*':ti,ab OR oncotypeDX*:ti,ab OR 21-gene:ti,ab OR gene-21:ti,ab OR 'DNA microarray kit'/exp OR mammaprint*:ti,ab OR mamma-print*:ti,ab OR 70-gene:ti,ab OR gene-70:ti,ab OR prosigna*:ti,ab OR 'microarray 50':ti,ab OR PAM50:ti,ab OR 'PAM 50':ti,ab OR 50-gene:ti,ab OR gene-50:ti,ab
Comparator	No search string
Outcomes	No search string
Limits	*Publication period* last 15 years (≥2009)
	No conference abstracts and preprints/select other publication types AND ([article]/lim OR [article in press]/lim OR [conference paper]/lim OR [conference review]/lim OR [data papers]/lim OR [editorial]/lim OR [erratum]/lim OR [letter]/lim OR [note]/lim OR [review]/lim OR [short survey]/lim)

Table 6: Cochrane Library

Population	[mh "Breast Neoplasms"] OR ("breast" NEXT neoplasm*):ti,ab OR ("breast" NEXT tumor*):ti,ab OR ("breast" NEXT tumour*):ti,ab OR ("breast" NEXT cancer):ti,ab OR ("breast" NEXT malignanc*):ti,ab OR ("breast" NEXT carcinoma*):ti,ab OR ("breast" NEXT adenocarcinoma*):ti,ab OR ("breast" NEXT sarcoma*):ti,ab OR ("mammary" NEXT neoplasm*):ti,ab OR ("mammary" NEXT tumor*):ti,ab OR ("mammary" NEXT tumour*):ti,ab OR ("mammary" NEXT cancer):ti,ab OR ("mammary" NEXT malignanc*):ti,ab OR ("mammary" NEXT carcinoma*):ti,ab OR ("mammary" NEXT adenocarcinoma*):ti,ab OR ("mammary" NEXT sarcoma*):ti,ab
Intervention	endopredict*:ti,ab OR EPclin:ti,ab OR "12-gene":ti,ab OR "gene-12":ti,ab OR ("oncotype" NEXT DX*):ti,ab OR oncotypeDX*:ti,ab OR "21-gene":ti,ab OR "gene-21":ti,ab OR mammaprint*:ti,ab OR mamma-print*:ti,ab OR "70-gene":ti,ab OR "gene-70":ti,ab OR prosigna*:ti,ab OR "microarray 50":ti,ab OR "PAM50":ti,ab OR "PAM 50":ti,ab OR "50-gene":ti,ab OR "gene-50":ti,ab
Comparator	No search string
Outcomes	No search string
Limits	*Publication period* last 15 years (≥2009)
	No conference abstracts and preprints NOT (congress:pt OR preprint:pt)

Table 7: ClinicalTrials.gov and EU Clinical Trials Register

Population	breast cancer
Intervention	Oncotype DX OR MammaPrint OR EndoPredict OR Prosigna
Comparator	No search string
Outcomes	No search string

C. Search strategy for economic evaluation systematic literature search

Table 8: PubMed (MEDLINE)

Population	"Breast Neoplasms"[Mesh] OR breast neoplasm*[tiab] OR breast tumor*[tiab] OR breast tumour*[tiab] OR breast cancer[tiab] OR breast malignanc*[tiab] OR breast carcinoma*[tiab] OR breast adenocarcinoma*[tiab] OR breast sarcoma*[tiab] OR mammary neoplasm*[tiab] OR mammary tumor*[tiab] OR mammary tumour*[tiab] OR mammary cancer[tiab] OR mammary malignanc*[tiab] OR mammary carcinoma*[tiab] OR mammary adenocarcinoma*[tiab] OR mammary sarcoma*[tiab]
Intervention	endopredict*[tiab] OR EPclin[tiab] OR 12-gene[tiab] OR gene-12[tiab] OR oncotype DX*[tiab] OR oncotypeDX*[tiab] OR 21-gene[tiab] OR gene-21[tiab] OR mammaprint*[tiab] OR mamma-print*[tiab] OR 70-gene[tiab] OR gene-70[tiab] OR prosigna*[tiab] OR microarray 50[tiab] OR PAM50[tiab] OR PAM 50[tiab] OR 50-gene[tiab] OR gene-50[tiab]
Comparator	No search string
Outcomes	“Technology Assessment, Biomedical”[Mesh] OR “Cost-Benefit Analysis”[Mesh] OR “Quality-Adjusted Life Years”[Mesh] OR technology assessment*[tiab] OR economic evaluat*[tiab] OR economic value[tiab] OR cost-benefit*[tiab] OR cost-effectiv*[tiab] OR cost-efficien*[tiab] OR cost-efficac*[tiab] OR cost-minim*[tiab] OR cost-utilit*[tiab] OR cost-consequen*[tiab] OR budget impact analys*[tiab] OR quality-adjusted life-year*[tiab] OR quality-adjusted lifeyear*[tiab] OR qaly*[tiab] [a]
Limits	*Publication period* last 15 years (≥2009)
	No conference abstracts and preprints NOT (congress[pt] OR preprint[pt])

Notes:
a = The economic search filter is a customised search filter for economic outcomes, which has been developed together with an information specialist. Existing economic search filters were used as input.

Table 9: Embase.com

Population	'breast tumor'/exp OR 'breast neoplasm*':ti,ab OR 'breast tumor*':ti,ab OR 'breast tumour*':ti,ab OR 'breast cancer':ti,ab OR 'breast malignanc*':ti,ab OR 'breast carcinoma*':ti,ab OR 'breast adenocarcinoma*':ti,ab OR 'breast sarcoma*':ti,ab OR 'mammary neoplasm*':ti,ab OR 'mammary tumor*':ti,ab OR 'mammary tumour*':ti,ab OR 'mammary cancer':ti,ab OR 'mammary malignanc*':ti,ab OR 'mammary carcinoma*':ti,ab OR 'mammary adenocarcinoma*':ti,ab OR 'mammary sarcoma*':ti,ab
Intervention	'breast cancer prognostic test kit'/exp OR endopredict*:ti,ab OR EPclin:ti,ab OR 12-gene:ti,ab OR gene-12:ti,ab OR 'oncotype DX*':ti,ab OR oncotypeDX*:ti,ab OR 21-gene:ti,ab OR gene-21:ti,ab OR 'DNA microarray kit'/exp OR mammaprint*:ti,ab OR mamma-print*:ti,ab OR 70-gene:ti,ab OR gene-70:ti,ab OR prosigna*:ti,ab OR 'microarray 50':ti,ab OR PAM50:ti,ab OR 'PAM 50':ti,ab OR 50-gene:ti,ab OR gene-50:ti,ab
Comparator	No search string
Outcomes	'biomedical technology assessment'/exp OR 'economic evaluation'/exp OR 'quality adjusted life year'/exp OR 'program cost effectiveness'/de OR ((technology NEAR/3 assessment*) OR (economic* NEAR/3 (evaluat* OR value)) OR ((cost OR costs) NEAR/3 (benefit* OR effectiv* OR efficien* OR efficac* OR minim* OR utilit* OR consequen*)) OR 'budget impact analys*':ti,ab OR (qualit* NEAR/3 adjust* NEAR/3 (life-year* OR lifeyear*)) OR qaly*):ab,ti [a]
Limits	*Publication period* last 15 years (≥2009)
	No conference abstracts and preprints AND ([article]/lim OR [article in press]/lim OR [conference paper]/lim OR [conference review]/lim OR [data papers]/lim OR [editorial]/lim OR [erratum]/lim OR [letter]/lim OR [note]/lim OR [review]/lim OR [short survey]/lim)

Notes:
a = The economic search filter is a customised search filter for economic outcomes, which has been developed together with an information specialist. Existing economic search filters were used as input.

Table 10: Cochrane Library

Population	[mh "Breast Neoplasms"] OR ("breast" NEXT neoplasm*):ti,ab OR ("breast" NEXT tumor*):ti,ab OR ("breast" NEXT tumour*):ti,ab OR ("breast" NEXT cancer):ti,ab OR ("breast" NEXT malignanc*):ti,ab OR ("breast" NEXT carcinoma*):ti,ab OR ("breast" NEXT adenocarcinoma*):ti,ab OR ("breast" NEXT sarcoma*):ti,ab OR ("mammary" NEXT neoplasm*):ti,ab OR ("mammary" NEXT tumor*):ti,ab OR ("mammary" NEXT tumour*):ti,ab OR ("mammary" NEXT cancer):ti,ab OR ("mammary" NEXT malignanc*):ti,ab OR ("mammary" NEXT carcinoma*):ti,ab OR ("mammary" NEXT adenocarcinoma*):ti,ab OR ("mammary" NEXT sarcoma*):ti,ab
Intervention	endopredict*:ti,ab OR EPclin:ti,ab OR "12-gene":ti,ab OR "gene-12":ti,ab OR ("oncotype" NEXT DX*):ti,ab OR oncotypeDX*:ti,ab OR "21-gene":ti,ab OR "gene-21":ti,ab OR mammaprint*:ti,ab OR mamma-print*:ti,ab OR "70-gene":ti,ab OR "gene-70":ti,ab OR prosigna*:ti,ab OR "microarray 50":ti,ab OR "PAM50":ti,ab OR "PAM 50":ti,ab OR "50-gene":ti,ab OR "gene-50":ti,ab
Comparator	No search string
Outcomes	[mh "Technology Assessment, Biomedical"] OR [mh "Cost-Benefit Analysis"] OR [mh "Quality-Adjusted Life Years"] OR technology assessment*:ti,ab OR economic evaluat*:ti,ab OR economic value:ti,ab OR cost-benefit*:ti,ab OR cost-effectiv*:ti,ab OR cost-efficien*:ti,ab OR cost-efficac*:ti,ab OR cost-minim*:ti,ab OR cost-utilit*:ti,ab OR cost-consequen*:ti,ab OR budget impact analys*:ti,ab OR quality-adjusted life-year*:ti,ab OR quality-adjusted lifeyear*:ti,ab OR qaly*:ti,ab [a]
Limits	*Publication period* last 15 years (≥2009)
	No conference abstracts and preprints NOT (congress:pt OR preprint:pt)

Notes:
a = The economic search filter is a customised search filter for economic outcomes, which has been developed together with an information specialist. Existing economic search filters were used as input.

Table 11: Tufts Medical Centre Cost-Effectiveness Analysis Registry and National Health Service Economic Evaluation Database

Population	keyword:("breast tumor" OR "breast neoplasm" OR "breast tumour" OR "breast cancer" OR "breast malignancy" OR "breast carcinoma" OR "breast adenocarcinoma" OR "breast sarcoma" OR "mammary neoplasm" OR "mammary tumor" OR "mammary tumour" OR "mammary cancer" OR "mammary malignancy" OR "mammary carcinoma" OR "mammary adenocarcinoma" OR "mammary sarcoma")
Intervention	keyword:("endopredict" OR "EPclin" OR "12-gene" OR "gene-12" OR "oncotype DX" OR "oncotypeDX" OR "21-gene" OR "gene-21" OR "mammaprint" OR "mamma-print" OR "70-gene" OR "gene-70" OR "prosigna" OR "microarray 50" OR "PAM50" OR "PAM 50" OR "50-gene" OR "gene-50")
Comparator	No search string
Outcomes	No search string
Limits	*Publication period* last 15 years (≥2009)